استناد به این مقاله: خسروی، عبدالرسول؛ فتاحی، رحمت الله؛ پریرخ، مهری؛ دیانی، محمدحسین (۱۳۹۲). بررسی کارآمدی کلیدواژه‌ها و عبارت‌های پیشنهادی موتور کاوش گوگل در بسط جستجو و افزایش ربط از دیدگاه دانشجویان تحصیلات تکمیلی. پژوهش‌نامه کتابداری و اطلاع‌رسانی،۳ (۱)، ۱۵۰-۱۳۳.

بررسی کارآمدی کلیدواژه‌ها و عبارت‌های پیشنهادی موتور کاوش گوگل در بسط جستجو و افزایش ربط از دیدگاه دانشجویان تحصیلات تکمیلی

عبدالرسول خسروی[۱]، دکتر رحمت‌الله فتاحی[۲]، دکتر مهری پریرخ[۳]، دکتر محمدحسین دیّانی[۴]

دریافت: ۱۳۹۰/۱۰/۱۲ پذیرش: ۱۳۹۱/۴/۴

چکیده

هدف: *هدف نظام‌های اطلاعاتی، بازیابی اطلاعات مرتبط با نیاز کاربر است. در این راستا، موتورهای کاوش قابلیّت جدیدی به نام "کلیدواژه‌ها و عبارت‌های پیشنهادی" پیاده نموده‌اند. این‌که این کلیدواژه‌ها و عبارت‌های پیشنهادی تا چه میزان با نیاز کاربران همخوان است، مسئله‌ای اساسی است. بر این اساس این پژوهش شکل گرفت تا کارآمدی کلیدواژه‌ها و عبارت‌های پیشنهادی موتور کاوش گوگل در بسط جستجو و افزایش ربط از دیدگاه کاربران را بررسی کند.*

روش: *این پژوهش کاربردی است و به روش پیمایشی انجام شده است. داده‌های مورد نیاز از طریق پرسش‌نامه به هنگام جستجوی کاربران در گوگل به دست آمد. بر این اساس، نمونه‌ای ۶۰ نفری از دو گروه علوم انسانی/اجتماعی و علوم پایه/فنّی- مهندسی در نظر گرفته شد و نشست‌هایی جداگانه با هر یک از آنها ترتیب داده شد.*

یافته‌ها: *یافته‌های پژوهش نشان داد که تفاوت معنی‌داری بین میزان ربط نتایج بازیافتی حاصل از عبارت جستجوی اولیه و نتایج بازیافتی حاصل از بسط جستجوی مبتنی بر کلیدواژه‌ها/ عبارت‌های پیشنهادی وجود دارد. یعنی کلیدواژه‌ها/ عبارت‌های پیشنهادی گوگل در بهبود ربط بازیافت‌ها مؤثر بوده است. همچنین تفاوت معنی‌داری از نظر نتایج بازیافتی حاصل استفاده از عبارت جستجوی اولیه و نتایج بازیافتی حاصل بسط جستجو بین دو حوزه مورد بررسی مشاهده نشد. در ادامه به منظور بهبود نتایج بازیافتی، توصیه‌های کاربردی و پژوهشی ارائه شده است.*

کلیدواژه‌ها: *بسط جستجو، موتورهای کاوش، گوگل، کلیدواژه‌ها و عبارت‌های پیشنهادی، ربط*

---

[1] عضو هیات علمی دانشگاه علوم پزشکی بوشهر Khosravi2422@gmail.com

[2] استاد گروه علم اطلاعات و دانش‌شناسی دانشگاه فردوسی

[3] استاد گروه علم اطلاعات و دانش‌شناسی دانشگاه فردوسی

[4] استاد گروه علم اطلاعات و دانش‌شناسی دانشگاه فردوسی

مقدمه

موتورهای کاوش به عنوان نظام بازیابی در ایجاد الگوهای مناسب جستجو با مشکلات عمده‌ای مواجه بوده‌اند (Ferreira & Atckinson, 2007) در عین حال، بسیاری از آنها با توسعه ویژگی‌های جستجو توانسته‌اند به کاربران در برآورده ساختن نیازهای اطلاعاتی آنها کمک نمایند. در حالی‌که برخی از این ویژگی‌ها توانسته است رضایت بیشتر کاربران از نتایج به دست آمده را حاصل کند، اما هنوز هم کاربران بسیاری، از نتایج حاصل از جستجوی مرتبط با نیازهای اطلاعاتی خود ناراضی هستند (Casasola & Gauch 1997; Chowdhury & Soboroff , 2002; Pokorny, 2004). البته این نارضایتی به دلیل وجود برخی از مشکلات بازیابی در موتورهای کاوش است؛ مانند دشواری در تدوین عبارت جستجو (Baeza-Yates., 2004; Lawrence & Giles, 1998; Schatz, Cochrane & Chen, 1996) دشواری در مرور صفحات بازیابی شده (Spink et al., 1998)؛ وجود الگوریتم‌های متفاوت جستجو (Ellis, Ford & Furner, 1998; Spink, et al., 1998) رویکردهای متفاوت کاربران در استفاده از کلیدواژه‌های واحد در جستجوهای خود (Sugiyama, Hatano & Yoshikawa 2004; Wang & Davison, 2008)، چگونگی اصلاح عبارت جستجو، دشواری در تدوین دوباره عبارت جستجو یا به اصطلاح «بسط جستجو» (Fattahi, Wilson & Cole, 2008; Jansen & Rieh, 2010; Spink, H. C. Ozmutlu, and S. Ozmutlu 2002; Widyantoro & Yen, 2001). پس از گذشت چند سال از طرّاحی موتورهای کاوش بسیاری از این مشکلات هنوز به قوت خود پا برجاست. پیشرفت‌های اخیر در زمینه هوشمندسازی محیط جستجو نشان می‌دهد که این محدودیت‌ها و مشکلات را می‌توان تا حدودی با استفاده از قابلیّت‌های جدید و نیز استفاده از دانش زمینه‌ای کاربر از پیش‌رو برداشت (Ferreira & Atckinson, 2007).

یکی از روش‌هایی که در شکل دادن رفتار اطلاع‌یابی کاربر مؤثّر است و به او کمک می‌کند تا رفتار اطلاع‌یابی خود را به خوبی مدیریت کرده و به نتیجه‌های بهتر برسد، ایجاد قابلیّت ارائه کلیدواژه‌ها و عبارت‌های پیشنهادی (برای بسط جستجو) در موتورهای کاوش است. از آنجایی که یافتن کلیدواژه‌ها و عبارت‌های مناسب برای بیان خواسته اطلاعاتی و به ویژه برای تدوین عبارت جستجوی مطلوب، همواره یکی از مشکلات اساسی کاربران در استفاده از نظام‌های اطلاعات از جمله موتورهای کاوش بوده است. برخی از این موتورها نظیر اسک[1]، گوگل[2]، یاهو[3] و ام‌اس‌ان[4] به ایجاد قابلیّت بسط جستجو با استفاده از

[1] www.Ask.com
[2] www.Google.com
[3] www.Yahoo.com
[4] www.MSN.com

کلیدواژه‌ها و عبارت‌های پیشنهادی مبادرت ورزیده‌اند. موتور کاوش گوگل از سال ۲۰۰۸ به منظور کمک به کاربران در تدوین عبارت جستجو، قابلیّت کلیدواژه‌ها و عبارت‌های پیشنهادی را به امکانات خود افزوده است. با توجّه به آنچه پیش‌تر در مورد مشکلات کاربران در جستجوی موفّق در موتورهای کاوش اشاره شد، این شیوه یعنی ارائه کلیدواژه‌ها یا عبارت‌های پیشنهادی بر اساس جستجوهای کاربران پیشین در موتور کاوش گوگل، می‌تواند یک راهبرد مناسب برای بسط جستجو در نظر گرفته شود. البته توسعه کارآمد چنین قابلیّت‌هایی نیازمند بررسی و پژوهش بیشتر است.

به طور کلی می‌توان گفت که ارائه کلیدواژه‌ها و عبارت‌های پیشنهادی، به عنوان یک قابلیت تازه بسط جستجو به منظور بازیابی اطلاعات مرتبط در موتورهای کاوش مورد توجّه قرار گرفته است. در زمان آغاز جستجو و هنگامی که کاربر شروع به درج کلمات در خانه جستجو می‌نماید، موتور کاوش در این باره که وی قصد جستجوی چه مفهومی را دارد، حدس می‌زند. از این‌رو، کلیدواژه‌ها و عبارت‌هایی را در آن رابطه پیشنهاد می‌کند. به بیان دیگر، در حین آن که کاربر کلمه مورد نظر خود را در خانه جستجو وارد می‌کند، موتور کاوش پیشنهادهای ممکن را در قالب کلیدواژه‌ها و عبارت‌های مرتبط، به وی نشان می‌دهد. بررسی‌های اولیه نشان می‌دهد که عواملی مانند: پی بردن به املای صحیح واژه، شناسایی ساختار صحیح عبارت که مفهوم مورد نظر را برساند، شناسایی جنبه‌های مختلف یک مفهوم، کاربرد یک عبارت در بافت مفهومی خاص که معنای کلیدواژه و یا عبارت را متفاوت می‌کند، می‌توانند در گزینش کلیدواژه‌ها و عبارت‌های پیشنهادی توسط کاربران مؤثر باشند.

آشکار است که در بسط جستجو با استفاده از کلیدواژه‌ها و عبارت‌های پیشنهادی در موتورهای کاوش، بین کلیدواژه و عبارت جستجوی کاربر با کلیدواژه‌ها و عبارت‌های پیشنهادی موتورهای کاوش نقطه اشتراکی وجود دارد. این امر موجب می‌شود تا کاربرانی که در تدوین عبارت جستجوی خود مشکل دارند، بیشتر از آن‌ها استفاده نمایند. اما مشاهده‌های پژوهشگر و بررسی‌های اوّلیه نشان داد که برخی از کاربران هنگام روبرو شدن با کلیدواژه‌ها و عبارت‌های پیشنهادی موتور کاوش گوگل نه تنها از آنها استفاده نمی‌کنند، بلکه به آنها توجّهی هم ندارند. به طور کلّی، با توجّه به اینکه رویکرد بسط جستجو مبتنی بر استفاده از کلیدواژه‌ها و عبارت‌های پیشنهادی بر اساس جستجوهای پیشین کاربران است، مشخص نیست که تا چه میزان این کلیدواژه‌ها و عبارت‌های پیشنهادی موتورهای کاوش از دیدگاه کاربران کارآمد هستند؟ و چنانچه کارآمد هستند، چگونه و تا چه میزان می‌توانند به کاربران در فرآیند رسیدن به اطمینان در مورد تدوین عبارت درست جستجو و بازیابی مدارک مرتبط‌تر کمک کنند؟

مروری کوتاه بر پیشینه پژوهش

بر اساس پژوهش‌های انجام شده، از دهه ۱۹۷۰، فنون بسط جستجو همواره مورد مطالعه پژوهشگران قرار گرفته است.

بررسی مدل‌های مورد استفاده در بسط جستجو نیز یکی از حوزه‌های مورد پژوهش بسط جستجو بوده است. در این راستا، کیی (Cui, 2003) در پژوهشی با عنوان "بسط جستجوی احتمالاتی با استفاده از سیاهه‌های عبارت جستجو" با هدف استخراج رابطه احتمالی بین کلیدواژه‌های جستجو و کلیدواژه‌های مدرک از طریق تحلیل سیاهه‌های داد و گرفت کاربران، شیوه جدیدی جهت بسط جستجوی مبتنی بر سیاهه‌های داد و گرفت پیشنهاد کرد و آن را مورد آزمون قرار داد. نتیجه بررسی‌ها نشان داد روش پیشنهادی وی کارکرد جستجو را بهبود بخشیده و باعث افزایش دقت در نتیجه‌های جستجو خواهد شد.

در حوزه اثربخشی شیوه‌های بسط جستجو، بیلربک و همکارانش (Billerbeck & Zobel., 2003) پژوهشی تحت عنوان «بسط جستجو با استفاده از عبارت‌های هم‌بسته[1]» با هدف بهبود اثربخشی بازیابی در موتورهای کاوش با استفاده از روش بسط جستجو انجام دادند. نتایج پژوهش آنان نشان دهنده بهبود ۲۶ تا ۲۹ درصدی نسبت به بازیابی غیر بسط یافته، یعنی استفاده نکردن از بسط جستجو و بهبود ۱۸ تا ۲۰ درصدی این روش نسبت به روش‌های قبلی بود.

یکی دیگر از رویکردهای پژوهشی، بررسی مشکلات استفاده از کلیدواژه‌ها و عبارت‌های پیشنهادی به منظور کمک به بسط جستجو بوده است. وانگ (Wang, 2008) در پژوهش خود به منظور بررسی مشکلات استفاده از کلیدواژه‌ها و عبارت‌های پیشنهادی، از نظام پیشنهاد برچسب[2] استفاده نمود. نتایج نشان داد که عبارت‌های پیشنهادی این نظام از دیدگاه کاربران به نسبت عبارت‌های پیشنهادی یاهو و گوگل از کیفیت بهتری برخوردارند. همچنین بین نتایج بازیابی شده از کلیدواژه‌ها و عبارت‌های پیشنهادی موتورهای کاوش یاهو و گوگل با نظام پیشنهادی برچسب، تفاوت معنی‌داری وجود دارد.

فتاحی و همکارانش (Fattahi, et al., 2008) به منظور شناسایی و تحلیل واژگان عمومی در منابع وب با رویکرد بسط جستجو و استفاده از زبان طبیعی در موتور کاوش گوگل پژوهشی به انجام رساندند. نتایج آن‌ها نشان داد که کاربرد واژگان عمومی (که همراه کلیدواژه‌های موضوعی می‌آیند) برای بسط جستجو می‌تواند نتایج مرتبط‌تر و دقیق‌تری در بر داشته باشد.

---

[1] Query Expansion Using Associated Queries
[2] Tag Suggestion System

در سال ۲۰۱۰ پژوهش‌های متعددی در حوزه بسط جستجو در کشورهای مختلف انجام شده است. سونگ و هی (Song & He, 2010) به منظور بررسی کارآمدی مدل پیشنهادی بسط جستجوی خویش پژوهشی تجربی و یک ماهه به انجام رساندند. به این منظور گزارش‌های سیاهه‌های داد و گرفت نشست-های جستجو از یک موتور کاوش تجاری بزرگ با حدود بیش از چهل میلیون عبارت جستجو مورد بررسی قرار گرفت. نتیجه این پژوهش نشان داد که مدل پیشنهادی آن‌ها در مقایسه با مدل‌های ربط کاذب و مدل‌های بسط جستجوی سنتی کارکرد بهتری دارند.

در داخل ایران و در حوزه بسط جستجو، پژوهشی نزدیک به موضوع مطالعه حاضر به انجام رسیده است. طاهریان و حیاتی (۱۳۸۶) در پژوهشی به منظور بررسی مقایسه‌ای میزان رضایت از الگوی جستجوی کلیدواژه‌ای بدون دریافت کمک و الگوی جستجوی کلیدواژه‌ای با دریافت کمک در موتور جستجوی گوگل از دیدگاه کاربران پرداخته‌اند. یافته‌های پژوهش نشان داد که حدود ۷۰٪ پیشنهادهای "گوگل ساجست بتا" به عنوان ارائه دهنده یک الگوی کلیدواژه‌ای با دریافت کمک از ربط متوسط به بالا برخوردار بوده است. همچنین بین الگوی جستجوی کلیدواژه‌ای بدون دریافت کمک و الگوی جستجوی کلیدواژه‌ای با دریافت کمک از لحاظ میزان دقت بر مبنای ربط کاربر مدار اختلاف معنی‌داری وجود نداشت، اما الگوی دوم منجر به افزایش معنی‌دار دقت بر مبنای ربط نظام‌مدار و میزان دقّت کلی گردیده است. در نهایت الگوی دوم موجب افزایش معنی‌دار میزان رضایت از سهولت استفاده و عملکرد کلی نظام در دستیابی به منابع مورد نظر گردید. لیکن این اختلاف بین میزان رضایت از زمان صرف شده در فرآیند جستجو و ارزش کلی نتایج جستجو معنی‌دار نبود. ولی رضایت کلی کاربران در روش دوم به طور معنی‌داری افزایش یافت.

بررسی متون مرور شده نشان می‌دهد که به دلیل نو بودن موضوع بسط جستجو ، اغلب پژوهش‌ها با روش‌های پیمایشی و تجربی و رویکرد کمّی و با استفاده از ابزارهایی نظیر پرسش‌نامه، مشاهده و سیاهه-های داد و گرفت انجام شده و جامعه مورد پژوهش اغلب آن‌ها کلیدواژه‌های پیشنهادی و متخصصان علوم رایانه بوده است. در حوزه بسط جستجو به زبان طبیعی، دیدگاه سایر کاربران به ویژه دانشجویان مورد بررسی قرار نگرفته است. علی‌رغم توجّه پژوهشگران به بسط جستجو در حوزه بازیابی اطلاعات و ارائه راه‌کارها و تکنیک‌های لازم و توجّه به هدف غایی که کاربر می‌باشد، اما دیدگاه کاربران به طور جدی مورد توجّه قرار نگرفته است. لازم است که بسط جستجو در محیط‌های وب به ویژه در موتورهای کاوش و در راستای چگونگی کارآمدی عبارت‌های پیشنهادی در بازیابی اطلاعات مرتبط از دیدگاه کاربران

مورد بررسی دقیق قرار گیرد.

فرضیه‌های پژوهش

۱. از دیدگاه کاربران، بین میزان ربط بازیافت‌های حاصل از عبارت جستجوی اولیه و بازیافت‌های حاصل از بسط جستجوی مبتنی بر کلیدواژه‌ها/ عبارت‌های پیشنهادی تفاوت وجود دارد.

۲. میان دیدگاه کاربران دو حوزه موضوعی علوم انسانی و اجتماعی و علوم پایه/ فنی- مهندسی در مورد میزان ربط بازیافت‌های حاصل از عبارت جستجوی اولیه و بازیافت‌های حاصل از بسط جستجوی مبتنی بر کلیدواژه‌ها/ عبارت‌های پیشنهادی تفاوت وجود دارد.

روش‌شناسی پژوهش

پژوهش حاضر از نوع کاربردی است، زیرا انتظار می‌رود از نتایج آن در جهت بهسازی الگوریتم‌های مناسب قابلیّت‌های موجود در موتورهای کاوش، طرّاحی کاربرمدارتر صفحه‌های رابط کاربر نظام‌ها و به طور خاص، توسعه قابلیّت ارائه کلیدواژه‌ها و عبارت‌های پیشنهادی استفاده شود. برای انجام این پژوهش از روش پیمایشی استفاده شد. و ابزار پرسش‌نامه جهت گردآوری داده‌ها مورد استفاده قرار گرفت. با توجه به این که پرسش‌نامه این پژوهش استاندارد نبوده است، جهت تعیین روایی، از نظر اساتید مجرب استفاده گردید (روایی صوری). همچنین برای کسب اطمینان بیشتر از صحیح بودن پرسش‌ها و رابطه آن‌ها با هدف از طراحی هر بخش (روایی محتوایی)، پرسش‌نامه در چند مرحله در اختیار افراد مختلف قرار گرفت و پس از گردآوری نظرات دوباره اصلاح شد. پایایی پرسش‌نامه مورد استفاده در این پژوهش به کمک آلفای کرونباخ ارزیابی گردید.

جامعه مورد پژوهش را دانشجویان مقاطع تحصیلات تکمیلی دانشگاه فردوسی مشهد که در سال تحصیلی ۹۰-۱۳۸۹ در مقطع تحصیلات تکمیلی به تحصیل اشتغال دارند، تشکیل می‌دهد. با توجه به رویکرد پژوهش، نمونه‌های پژوهش از انواع نمونه‌گیری غیر احتمالی و از نوع نمونه هدفمند انتخاب شد (پاول، ۱۳۷۹ ص ۱۱۰). بدین منظور، یک نمونه ۶۰ نفری داوطلب برای دو گروه از حوزه‌های مورد مطالعه (علوم انسانی/ اجتماعی و علوم پایه/ فنّی- مهندسی) در نظر گرفته شد. بدیهی است به منظور انتخاب ۶۰ نفر واجد شرایط انجام پژوهش یعنی، آشنا به جستجو در موتور کاوش گوگل، آشنا با زبان انگلیسی و استفاده از بسط جستجو در هر حوزه، روش‌های معمول در نمونه‌گیری تصادفی کارایی چندانی نخواهند داشت. در چنین شرایطی استفاده از انواع روش‌های نمونه‌گیری غیرتصادفی پیشنهاد می‌شود. در

این پژوهش به منظور بررسی موضوع در دو حوزه پیش‌گفته، نشست‌هایی جداگانه با هر یک از ۶۰ نفر ترتیب داده شد.

یافته‌های پژوهش

از دیدگاه کاربران، بین میزان ربط نتایج بازیابی حاصل از عبارت جستجوی اولیه و نتایج بازیابی حاصل از بسط جستجوی مبتنی بر کلیدواژه‌ها/ عبارت‌های پیشنهادی تفاوت وجود دارد.

بسط جستجو به منظور تدوین عبارت جستجو به شکل مناسب‌تر انجام می‌شود و هدف آن افزایش میزان ربط در نتایج بازیابی است، به صورتی که منابع بازیابی شده جنبه‌ها یا ویژگی مورد نظر جستجوگر را در برداشته باشد، این جنبه‌ها می‌تواند شامل سطح مطلب، مخاطبان مورد نظر، عمق موضوع، رویکرد یا نوع منبع باشد.

برای بررسی مقایسه و تأثیر استفاده از کلیدواژه‌ها/ عبارت‌های پیشنهادی موتور کاوش گوگل در میزان ربط بازیافت‌ها، در ابتدا از کاربران خواسته شد تا با استفاده از کلیدواژه و عبارت اولیه خود در زمینه موضوع مورد علاقه و مرتبط با پژوهش (بدون انجام بسط جستجو)، جستجویی انجام به انجام رسانند و هر یک از آنها ۱۰ بازیافت (صفحه‌های بازیابی شده) اولیه موتور کاوش گوگل را ارزیابی کنند. این بازیافت-ها در ستون اول جدول زیر نشان داده شده است. در هر ردیف از این ستون هر یک از ۶۰ کاربر می‌بایست یک بازیافت را بررسی کند. به این ترتیب ۶۰ بازیافت مورد بررسی قرار گرفت. سپس از آنها خواسته شد تا میزان ارتباط هر یک از بازیافت‌ها را در پرسش‌نامه بر اساس طیف‌های تعیین شده، مشخص نمایند. نتایج بررسی پاسخ‌ها در مورد میزان ربط بازیافت‌ها در جدول شماره ۱ نشان داده شده است.

جدول ۱: فراوانی و درصد فراوان میزان مرتبط بودن بازیافت‌ها در جستجوی اولیه (پیش از بسط جستجو (n=۶۰)

بازیافت‌ها		**طیف ارزیابی**				میانگین	انحراف معیار
		ناکافی[1]	بدون ارتباط	کم ربط	مرتبط		
بازیافت -۱	فراوانی	۷	۷	۵	۴۱	۲/۳۳	۱/۰۸۴
	درصد	۱۱/۷	۱۱/۷	۸/۳	۶۸/۳		
بازیافت -۲	فراوانی	۶	۹	۷	۳۸	۲/۲۸	۱/۰۵۹
	درصد	۱۰	۱۵	۱۱/۷	۶۳/۳		
بازیافت -۳	فراوانی	۱۰	۷	۱۸	۲۴	۱/۹۵	۱/۱۰۵
	درصد	۱۶/۷	۱۱/۷	۳۰	۴۰		
بازیافت -۴	فراوانی	۱۰	۸	۲۱	۲۱	۱/۸۸	۱/۰۷۵
	درصد	۱۶/۷	۱۳/۳	۳۵	۳۵		
بازیافت -۵	فراوانی	۱۱	۶	۹	۳۲	۲/۰۷	۱/۱۹۷
	درصد	۱۸/۳	۱۰	۱۵	۵۳/۳		
بازیافت -۶	فراوانی	۹	۵	۱۵	۳۰	۲/۱۲	۱/۱۰۰
	درصد	۱۵	۸/۳	۲۵	۵۰		
بازیافت -۷	فراوانی	۷	۳	۱۴	۳۴	۲/۲۹	۱/۰۲۶
	درصد	۱۱/۷	۵	۲۳/۳	۵۶/۷		
بازیافت -۸	فراوانی	۱۳	۸	۱۳	۲۶	۱/۸۷	۱/۲۰۰
	درصد	۲۱/۷	۱۳/۳	۲۱/۷	۴۳/۳		
بازیافت -۹	فراوانی	۱۱	۶	۱۷	۲۵	۱/۹۵	۱/۱۳۶
	درصد	۱۸/۳	۱۰	۲۸/۳	۴۱/۷		
بازیافت -۱۰	فراوانی	۶	۴	۱۷	۲۹	۲/۲۳	۰/۹۹۱
	درصد	۱۰	۶/۷	۲۸/۳	۴۸/۳		

بر پایه اطلاعات جدول ۱، از ده بازیافت (سایت یا صفحه بازیابی شده) که در صفحه اول گوگل و بر اساس جستجوی کلیدواژه یا عبارت اولیه نمایان شده بود، ۶ مورد مرتبط‌تر تشخیص داده شد. به بیانی دیگر، با نگاهی دقیق‌تر به جدول ۱ در می‌یابیم که میانگین میزان ربط شش مورد بازیابی شده در صفحه اول موتور کاوش گوگل، در مقایسه با حداکثر میانگین، بالاتر از ۱/۵ (نقطه برش) بوده است. البته با توجه به نمره انحراف معیار بدست آمده در هر مورد می‌توان گفت که نمره انحراف بالا در این جدول نشان می‌-دهد که توافق کاربران بر مرتبط بودن ده نتیجه بازیابی شده صفحه نخست در چهار گزینه داده شده

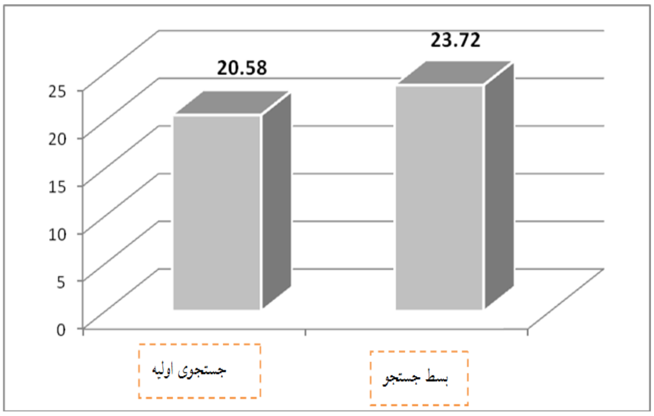
(مرتبط، کم ربط، ناکافی و بدون ارتباط) یکسان نیست، یعنی نمی‌توان نتیجه گرفت که میانگین نسبی میزان ربط نتیجه‌های بازیابی شده صفحه اول موتور کاوش گوگل، در مورد اغلب جستجوها مرتبط یا بدون ارتباط هستند، زیرا پراکندگی زیادی در بین پاسخ‌های کاربران دیده می‌شود.

در مرحله بعد، از شرکت‌کنندگان خواسته شد تا با گزینش یک مورد از کلیدواژه‌ها/ عبارت‌های پیشنهادی موتور کاوش که از دیدگاه آنها و به لحاظ مفهومی به عبارت جستجوی اولیه آنها نزدیک‌تر است، دوباره به جستجو بپردازند. نتایج بررسی پاسخ‌های ۶۰ آزمودنی در مورد میزان ربط ده بازیافت نخست، در جدول ۲ نشان داده شده است.

جدول ۲. فراوانی میزان مرتبط بودن ده بازیافت پس از بسط جستجو (n=۶۰)

نتیجه‌های بازیابی شده		**طیف ارزیابی**				میانگین	انحراف معیار
		ناکافی	بدون ارتباط	کم ربط	مرتبط		
بازیافت شماره ۱-	فراوانی	۰	۶	۱۰	۴۴	۲/۵۳	۰/۹۲۹
	درصد	۰	۱۰	۱۶/۷	۷۳/۳		
بازیافت شماره ۲-	فراوانی	۶	۴	۱۰	۴۰	۲/۴۳	۰/۹۲۷
	درصد	۱۰	۶/۷	۱۶/۷	۶۶/۷		
بازیافت شماره ۳-	فراوانی	۲	۵	۱۲	۴۰	۲/۴۷	۰/۹۱۶
	درصد	۳/۳	۸/۳	۲۰	۶۶/۷		
بازیافت شماره ۴-	فراوانی	۳	۳	۱۳	۴۱	۲/۵۳	۰/۸۱۲
	درصد	۵	۵	۲۱/۷	۶۸/۳		
بازیافت شماره ۵-	فراوانی	۳	۹	۱۰	۳۷	۲/۲۷	۱/۱۱۲
	درصد	۵	۱۵	۱۶/۷	۶۱/۷		
بازیافت شماره ۶-	فراوانی	۲	۹	۱۲	۳۶	۲/۲۷	۱/۰۹۶
	درصد	۳/۳	۱۵	۲۰	۶۰		
بازیافت شماره ۷-	فراوانی	۲	۴	۱۱	۴۳	۲/۵۵	۰/۸۵۲
	درصد	۳/۳	۶/۷	۱۸/۳	۷۱/۷		
بازیافت شماره ۸-	فراوانی	۳	۵	۱۱	۴۰	۲/۴۶	۰/۹۳۴
	درصد	۵	۸/۳	۱۸/۳	۶۶/۷		
بازیافت شماره ۹-	فراوانی	۳	۸	۱۴	۳۵	۲/۲۷	۱/۰۵۶
	درصد	۵	۱۳/۳	۲۳/۳	۵۸/۳		
بازیافت شماره ۱۰-	فراوانی	۳	۱۲	۱۳	۳۲	۲/۰۸	۱/۱۸۳
	درصد	۵	۲۰	۲۱/۷	۵۳/۳		

با نگاهی به جدول ۲ مشاهده می‌شود که نمره میانگین هر ده بازیافت در صفحه اول گوگل که کاربران با استفاده از کلیدواژه‌ها/ عبارت‌های پیشنهادی موتور کاوش گوگل به دست آورده‌اند، بالاتر از ۱/۵ (در مقایسه با حداکثر میانگین ۳) بوده است، به بیانی، میزان ربط این بازیافت‌ها در مقایسه با ده بازیافت پیشین (بدون استفاده از کلیدواژه‌ها/ عبارت‌های پیشنهادی کاربران (جدول۲))، بیشتر بوده است. با توجه به نمره انحراف معیار بدست آمده در مورد هر نتیجه می‌توان گفت که توافق کاربران بر مرتبط بودن ده نتیجه بازیابی شده صفحه نخست در چهار طیف داده شده (طیف مرتبط، کم ربط، ناکافی و بدون ارتباط) به طور نسبی بهتر از مورد قبل بوده است (نمودار شماره ۱). یعنی، می‌توان نتیجه گرفت میانگین ربط بازیافت‌ها پس از استفاده از کلیدواژه‌ها/ عبارت‌های پیشنهادی (بسط جستجو) بالاتر از میانگین بازیافت‌های حاصل جستجوی کلیدواژه یا عبارت اولیه کاربر است. در عین حال، نمی‌توان گفت که این نسبت در همه نتایج یکسان است.



نمودار ۱: میانگین نظرات کاربران در مورد میزان ربط بازیافت‌های حاصل از جستجوی اولیه و جستجوی بسط یافته (مبتنی بر کلیدواژه/ عبارت‌های پیشنهادی)

با توجه به این که نتیجه‌های جستجوی مبتنی بر استفاده از کلیدواژه‌ها و عبارت‌های پیشنهادی دارای توزیع غیرنرمال می‌باشد، بنابراین از آزمون ناپارامتری ویلکاکسون برای آزمون فرضیه وجود تفاوت بین میزان ربط نتایج بازیابی حاصل از عبارت جستجوی اولیه و نتایج بازیابی حاصل از بسط جستجوی مبتنی بر کلیدواژه‌ها/ عبارت‌های پیشنهادی استفاده شد.

با توجه به غیر نرمال بودن متغیّر جستجوی مبتنی بر کلیدواژه‌ها و عبارت‌های، بایستی فرض‌های زیر

مورد آزمون قرار گیرد:

فرض H_0 به معنای عدم وجود تفاوت بین میانگین میزان ارتباط صفحه‌های بازیابی حاصل از جستجوی اولیه و میانگین میزان ارتباط صفحه‌های بازیابی حاصل از کاربرد کلیدواژه‌ها است و فرض H_1 به معنای وجود تفاوت بین آن میانگین‌ها می‌باشد.

جدول۳: نتایج آزمون ویلکاکسون برای سنجش تفاوت معنی‌داری بین قضاوت ربط در مورد بازیافت‌ها دو مرحله جستجو (n=۶۰)

ردیف	آماره	مقدار
۱	آماره ویلکاکسون	-۳/۷۴۶
۲	سطح معنی‌داری	۰/۰۰۰

چون سطح معنی‌داری حاصل از آزمون ویلکاکسون برابر ۰/۰۰۰ می‌باشد، بنابراین فرض صفر رد و فرض یک پذیرفته می‌شود. یعنی بین میزان ربط بازیافت‌های حاصل از عبارت جستجوی اولیه و بازیافت-های حاصل از بسط جستجوی مبتنی بر کلیدواژه‌ها/ عبارت‌های پیشنهادی تفاوت معنی‌داری وجود دارد. به عبارتی، از نظر کاربران، کلیدواژه‌ها/ عبارت‌های پیشنهادی در بهبود ربط بازیافت‌ها مؤثر بوده است. با نگاهی به مجموع یافته‌ها، می‌توان به این نتیجه رسید که تفاوت معنی‌داری بین میزان ربط بازیافت‌های حاصل از کاربرد کلیدواژه‌ها/ عبارت‌های پیشنهادی موتور کاوش گوگل با بازیافت‌های حاصل از کاربرد کلیدواژه اولیه توسط کاربر وجود دارد. بنابراین، فرض وجود تفاوت بین آنها تأیید می‌شود. این نتیجه می‌تواند نشانه تأثیر مثبت کلیدواژه‌ها/ عبارت‌های پیشنهادی موتور کاوش گوگل بر میزان ربط بازیافت‌ها باشد.

فرضیه دوم: میان دیدگاه کاربران دو حوزه موضوعی علوم انسانی/اجتماعی و علوم پایه/ فنّی- مهندسی از نظر میزان ربط بازیافت‌های حاصل از عبارت جستجوی اولیه و بازیافت‌های حاصل از بسط جستجوی مبتنی بر کلیدواژه‌ها/ عبارت‌های پیشنهادی تفاوت وجود دارد.

برای مقایسه میزان ارتباط بازیافت‌ها با نیاز کاربران در دو حوزه موضوعی علوم انسانی و اجتماعی و علوم پایه و فنی/ مهندسی از آزمون t دو نمونه‌ای مستقل استفاده شد. فرض H_0 به معنای نبودن تفاوت بین میانگین میزان ارتباط بازیافت‌ها با نیاز کاربران در حوزه‌های مورد مطالعه و فرض H_1 به معنای وجود تفاوت بین میانگین میزان ارتباط بازیافت‌ها با نیاز کاربران در حوزه‌های مورد مطالعه می‌باشد.

جدول ۴: آماره‌های توصیفی متغیّر میزان ارتباط بازیافت‌ها با نیاز کاربران (n=۶۰)

حوزه	میانگین	انحراف معیار
علوم انسانی و اجتماعی	۲۰/۲۰	۶/۵۰۴
علوم پایه و فنی/ مهندسی	۲۰/۹۷	۶/۷۷۰

قبل از انجام آزمون t لازم است فرض برابری واریانس‌ها در دو گروه علوم انسانی و اجتماعی و علوم پایه و فنی/ مهندسی با آزمون Leven Test بررسی گردد. در صورتی که فرض برابری واریانس‌ها رد شود، می‌توان از فرمول‌های دیگری برای انجام این آزمون استفاده کرد.

جدول ۵: نتایج آزمون فرض برابری واریانس‌ها برای متغیّر میزان ارتباط بازیافت‌ها با نیاز کاربر (n=۶۰)

آماره‌های آزمون	مقدار
آماره فیشر F	۰/۴۷۰
سطح معنی‌داری	۰/۴۹۶

چون سطح معنی‌داری معادل ۰/۴۹۶ می‌باشد، بنابراین در سطح $\alpha = 0.05$ فرض برابری واریانس‌ها را می‌پذیریم.

جدول ۶: نتایج آزمون t مستقل برای متغیّر ارتباط بازیافت‌ها با نیاز کاربر (n=۶۰)

آماره‌های آزمون	مقدار
آماره t	-۰/۴۴۷
درجه آزادی	۵۸
سطح معنی‌داری	۰/۶۵۶

با توجه به نتایج حاصل از جدول ۶ چون سطح معنی‌داری معادل ۰/۶۵۶ می‌باشد، بنابراین در سطح $\alpha = 0.05$ فرض صفر را رد نمی‌شود.

نتیجه‌گیری

پژوهش پیرامون بسط جستجو با استفاده از کلیدواژه‌ها و عبارت‌های پیشنهادی می‌تواند به طراحی و توسعه نظام‌های بازیابی اطلاعات بهتر با اثربخشی بیشتر منجر شود. طراحان موتورهای کاوش با فراهم نمودن رابط کاربری مناسب‌تر بر اساس دیدگاه‌ها و رفتارهای کاربران، می‌توانند نظام‌های خود را توسعه بخشند. قابلیّت ارائه کلیدواژه‌ها و عبارت‌های پیشنهادی موتورهای کاوش باید به گونه‌ای طراحی شود که نه تنها عملکرد بهتر کاربر در استفاده از موتور کاوش را موجب شود، بلکه به وی کمک کند تا درک بهتری از چنین محیطی به دست آورد. در طراحی نظام‌های کاوش، توجّه بیشتر به ارائه درست کلیدواژه‌ها و عبارت‌های پیشنهادی از این رو اهمیّت دارد که می‌تواند به طراحی کاربرمدارتر صفحات، عناصر و اجزاء اطلاعاتی، رابط‌های کاربر در موتورهای کاوش منجر شود و فرآیند جستجو را بهبود بخشد.

بنابراین، با توجّه به نتایج این پژوهش که نشان داد، کلیدواژه‌ها و عبارت‌های پیشنهادی توانسته است به کاربران در بازیابی نتایج مرتبط با نیازهای آنها کمک کند، توجّه به مسایل مرتبط با ویژگی‌های شناختی کاربران دارای اهمیّت زیادی است. برای جستجوگرانی که با مشکلاتی مانند آگاهی نداشتن از املاء صحیح واژه‌ها روبرو باشند و یا از کاربرد درست کلیدواژه‌ها در عبارت جستجو به صورتی که مفهوم مورد نظر آنان را برساند اطلاع نداشته باشند و یا ندانند چه جنبه‌ای از یک موضوع مورد نظر آنهاست، مشاهده سیاهه‌ای از کلیدواژه‌ها و عبارت‌های پیشنهادی راه‌گشاست تا چنین مشکلاتی برطرف شود و یافتن کلیدواژه‌ها و یا عبارت‌های مورد نظر آنان سهولت یابد. طراحان موتورهای کاوش، با تمهیداتی مناسب در طراحی صفحه‌های رابط کاربر می‌توانند ضمن استفاده از دانش زمینه‌ای کاربران در استخراج کلیدواژه‌ها و عبارت‌های پیشنهادی در ارائه مناسب آنها به کاربران کمک کند. زیرا این موارد به آنها کمک می‌کند تا بصورت منطقی و کنترل شده به بسط عبارت جستجوی خود بپردازد (Back, 2001).

توصیه‌های کاربردی و پیشنهادهایی برای پژوهش‌های آینده

در پیوند با یافته‌های پژوهش حاضر پیشنهادهای زیر در راستای بهبود وضعیت طراحی موتورهای کاوش و قابلیّت‌های آن به ویژه چگونگی ارائه کلیدواژه‌ها و عبارت‌های پیشنهادی (بسط جستجو) عرضه می‌گردد:

۱. در راستای یادگیری بهتر و نحوه تعامل کارآمدتر کاربران با موتورهای کاوش به ویژه آشنایی آنها با قابلیّت‌های جستجو و بسط جستجو، پیشنهاد می‌شود کارگاه‌های آموزشی متناسب با گروه‌های مختلف کاربران از سوی کتابخانه‌ها برگزار شود.

۲. با توجه به کارآمدی کلیدواژه‌ها و عبارت‌های پیشنهادی در تدوین درست عبارت جستجو و کمک به کاربران در بازیابی اطلاعات مرتبط‌تر، بر اساس نتایج این پژوهش، به طراحان نظام‌های بازیابی اطلاعات از جمله نرم‌افزارهای کتابخانه‌ای و پایگاه‌های اطلاعاتی (مقاله‌ها، پایان نامه‌ها) که قابلیّت جستجوی کلیدواژه‌ای دارند، توصیه می‌گردد که بنا به ضرورت در بکارگیری این قابلیّت در رابط کاربر نرم‌افزارهای خود اقدام نمایند.

۳. برخی از موتورهای کاوش نظیر "یاهو"، "اسک" و "بینگ" نیز قابلیّت ارائه کلیدواژه‌ها و عبارت‌های پیشنهادی را به امکانات خود اضافه نموده‌اند. به نظر می‌رسد که برخی از آنها از جمله یاهو کلیدواژه‌ها و عبارت‌های پیشنهادی را با ساختار متفاوتی ارائه می‌دهند. از این‌رو، پیشنهاد می‌شود که پژوهشی مقایسه‌ای درباره کارآمدی و ساختار کلیدواژه‌ها و عبارت‌های پیشنهادی و نتایج بازیافتی در هر یک از این موتورها به منظور تعیین شباهت‌ها و تفاوت‌های آن‌ها انجام شود. نتایج آن پژوهش می‌تواند در تعیین ساختار و الگوریتم مناسب به طراحان، و در گزینش بهتر ابزار و راهبرد جستجو به کاربران یاری رساند.

۴. این پژوهش به کارآمدی کلیدواژه‌ها و عبارت‌های پیشنهادی (برای بسط جستجو) از دیدگاه کاربران دو حوزه موضوعی علوم انسانی/ اجتماعی و علوم پایه/ فنّی – مهندسی پرداخته است. به منظور کسب اطمینان درباره تعمیم یافته‌های این پژوهش، توصیه می‌شود که پژوهش‌های مشابه دیگری با گروه‌های مختلف کاربران و در موضوع‌ها و شرایط دیگر انجام گیرد. چنین پژوهش‌هایی می‌تواند در شناسایی عناصر و عوامل مشترک مورد توجه گروه‌های کاربران و چگونگی استفاده از این قابلیّت و کارآمدی بیشتر آن سودمند باشد.

کتابنامه

پاول، رونالد (۱۳۷۹). روش‌های اساسی پژوهش برای کتابداران. (نجلا حریری، مترجم). تهران: مرکز علمی انتشارات دانشگاه آزاد اسلامی.

طاهریان، آمنه؛ حیاتی، زهیر (۱۳۸۸). بررسی مقایسه‌ای الگوی جستجوی کلیدواژه‌ای با دریافت کمک و الگوی جستجوی بدون دریافت کمک در موتور کاوش گوگل از دیدگاه کاربران. *فصلنامه مطالعات تربیتی و روان‌شناسی دانشگاه فردوسی مشهد*، ۱۰ (۲). ۹۱-۱۱۲.

فتاحی، رحمت‌الله (۱۳۸۵). شناسایی و تحلیل واژگان عمومی در منابع وب: رویکردی نو به بسط عبارت جستجو با استفاده از زبان طبیعی در موتورهای کاوش. دو *فصلنامه مطالعات علوم تربیتی و روانشناسی*، ۷ (۱)، ۳۱-۵۳.

Baeza-Yates R. (2004). *Query Recommendation Using Query Logs in Search Engines.* 588--596.

Billerbeck, B., & Zobel, J. (2004). *Questioning Query Expansion: An Examination of Behaviour and Parameters*. New oreland, USA

Casasola, E., & Gauch, S. (1997). *Intelligent Information Agents for the World Wide Web*: Information and Telecommunication Technology Center, the University of Kansas.

Chowdhury, G., Soboroff, S. (2002). Automatic Evaluation of World Wide Web Search Services. Retieved from:
http://citeseer.ist.psu.edu/chowdhury02automatic.html

Ellis, D., Ford, N., Furner, J. (1998). In Search of the Unknown User: Indexing and Hypertext and the World Wide Web. *Journal of Documentation 54*(1), 28–47.

Fattahi, R., Wilson, CS, Cole, F (2008). An Alternative Approach to Natural Language Query Expansion in Search Engines Text Analysis of Non-topical Terms in Web Documents. *Information Processing and Management, 44*, 1503-1516.

Ferreira, A., Atkinson, J. (2007). Intelligent Search Engines. Retrieved from:
http://books.mcgraw-hill.com/EST10/site/supparticles/Intelligent-search-engines-YB070360.pdf

Giles, L. S. a. C. L. (1998). Context and page Analysis for Improved Web Search. *IEEE Internet Computing 2*(4).

Jansen, B., young Rieh, S. (2010). The Seventeen Theoretical Constructs of Information Searching and Information Retrieval. *Journal of the American Society for Information Science and Technology*, *61*(8).

Lawrence, S., & Giles, C. L. (1998). Context and page Analysis for Improved Web Search. *Internet Computing, IEEE, 2*(4), 38-46.

Natsev, A., Haubold, A., Tesic, J., Xie, L., & Yan, R. (2007). *Semantic Concept-Based Query Expansion and Re-Ranking for Multimedia Retrieval.*

Pokorny, J. (2004). Web Searching and Information Retrieval. *IEEE Computer Software, 6*(4), . 43–48

Qiu, Y. &. H. P. Frei. (1993). Concept Based Query Expansion. Retrieved from http://doi.acm.org/10.1145/160688.160713

Schatz, B, P. A. Cochrane, & H Chen (1996). Interactive Term Suggestion for Users of Digital Libraries: Using Subject Thesauri and Co-Occurrence lists for Information Retrieval. *Paper presented at the 1st ACM International Conference on Digital Libraries.*

Song, Y., & L, He. (2010). Optimal rare query suggestion with implicit user feedback.

Spink, A., H. C. Ozmutlu, and S. Ozmutlu (2002). Multitasking Information Seeking and Searching processes. *Journal of the American society for information science and technology 53*(8)): 639-652.

Spink, et. al. (1998). From Highly Relevant to not Relevant: Examining Different Regions of Relevance. In Information Processing and Management. Retrieved from: http://www.informatik.uni-trier.de/ley/db/indices/a-tree/s/Spink:Amanda.html

http://ir.mathcs.emory.edu/SSM2008/papers/ssm14-strohmaier.pdf

Sugiyama, K., Hatano, K. and Yoshikawa, M. (2004). *Adaptive Web Search Based on User Profile Constructed without any Effort from User.* Paper presented at the In Proceedings of WWW '04.

Wang, J., Davison, B. D. (2008). Explorations in Tag Suggestion and Query Expansion. Paper presented at the SSM '08: Proceeding of the 2008 ACM Workshop on Search in Social Media, pages 43–50, New York, NY, USA, ACM.

Widyantoro, D. H., Yen, J. (2001). A Fuzzy Ontology-based Abstract Search Engine and its User Studies. *Paper presented at the Proceedings of the 10th IEEE International Conference on.*